Average Nucleotide Identity (ANI)

Introduction

Average Nucleotide Identity (ANI) is a measure of nucleotide-level genomic similarity between the coding regions of two genomes. First, BBHs between a genome pair are computed as pairwise bidirectional best nSimScan hits of genes having 70% or more identity and at least 70% coverage of the shorter gene.

Alignment Fraction (AF) is computed using to the following formula:

$$AF_{G1 \rightarrow G2} = \frac{\text{lengths of BBH genes}}{\sum \text{length of genes in genome 1}}$$

Average Nucleotide Identity (ANI) is computed using the following formula:

$$gANI_{G1 \rightarrow G2} = \frac{\sum_{bbh} (\text{Percent Identity} * \text{Alignment length})}{\text{lengths of BBH genes}}$$

Pairwise ANI

BBHs between a genome pair are computed as pairwise bidirectional best nSimScan hits of genes having 70% or more identity and at least 70% coverage of the shorter gene. Genome(s) can either be selected from IMG or uploaded as a nucleotide sequence in FASTA format (using the Upload File button) to compute ANI to selected genome(s) in IMG.

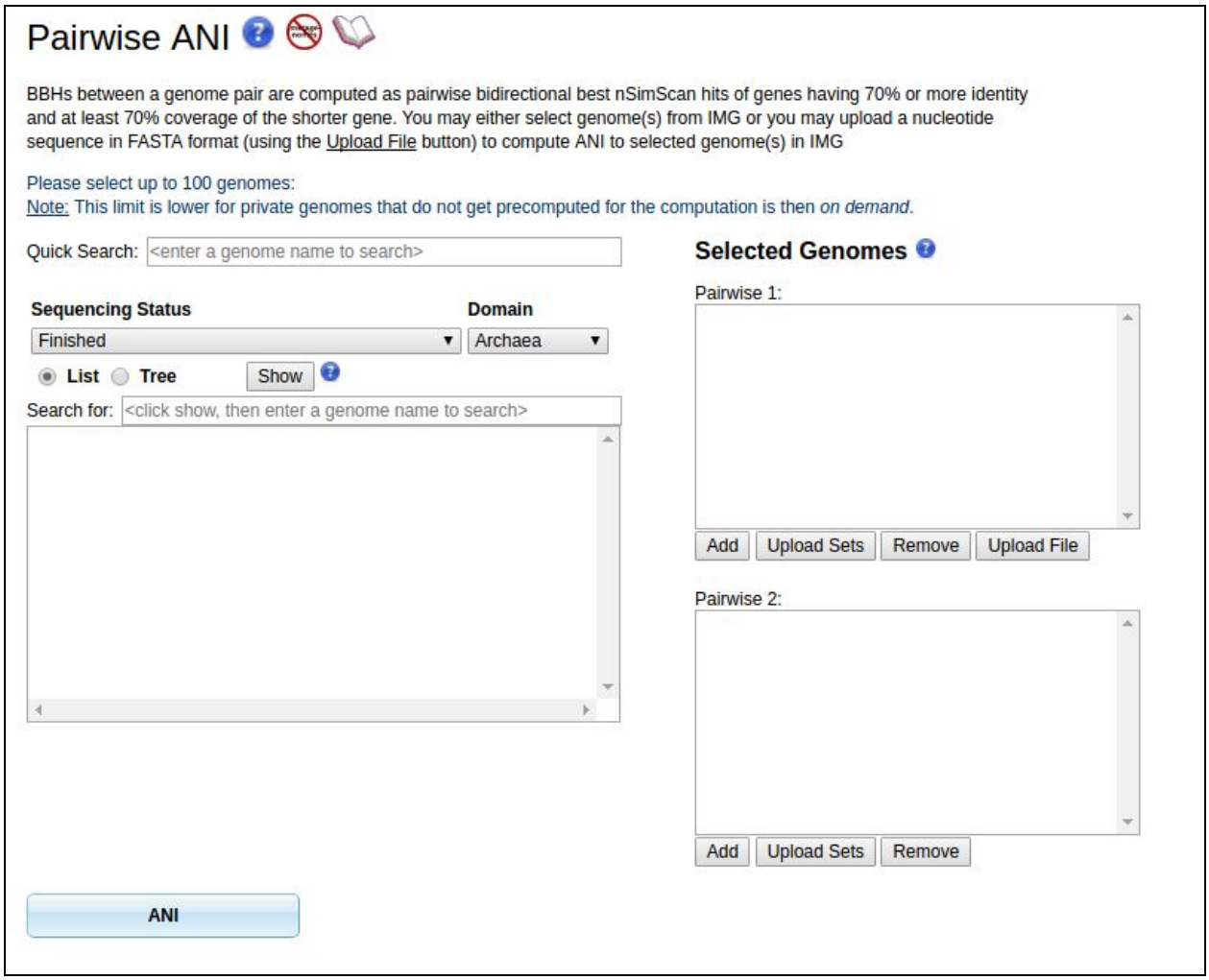
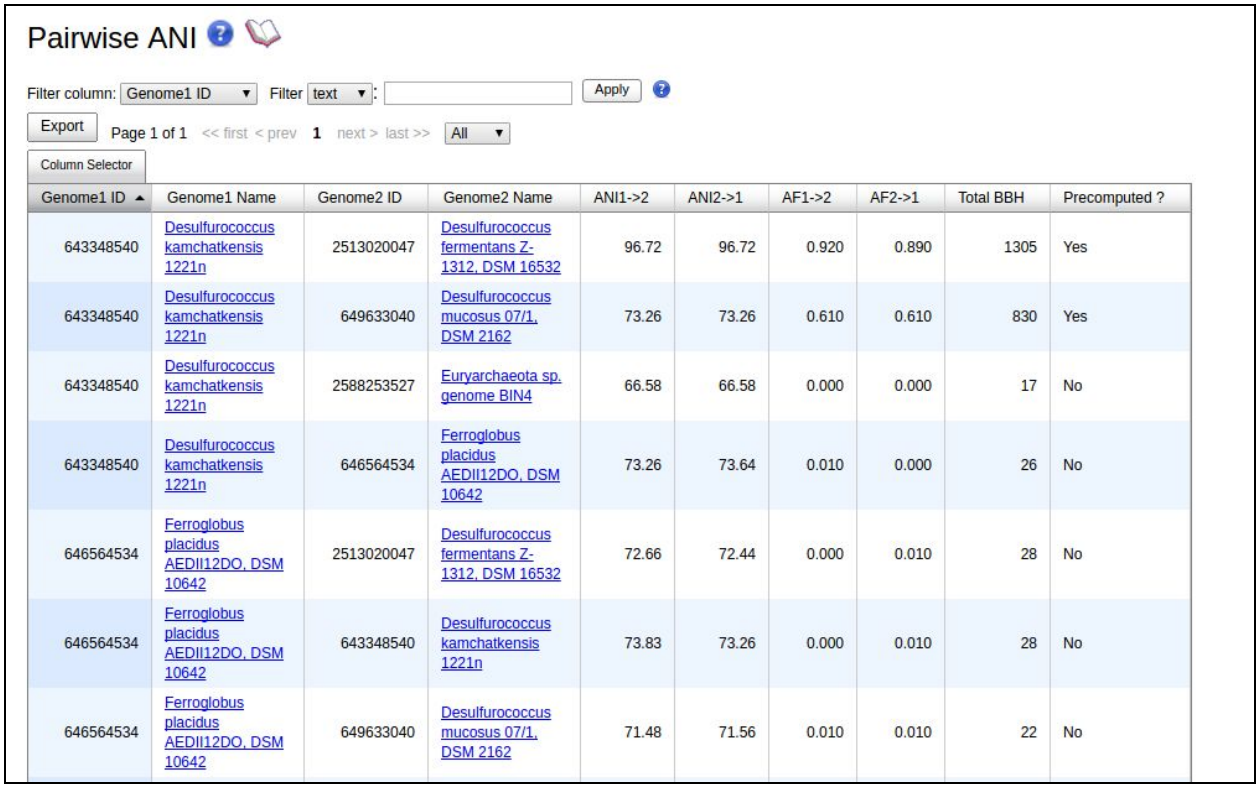


Figure 1. ANI Pairwise

Pairwise ANI is precomputed for most genomes. Private genomes and newly added genomes will be precomputed on-demand which will slow down the computation (Fig. 1b).



Pairwise ANI

Filter column: Genome1 ID Filter text : Apply

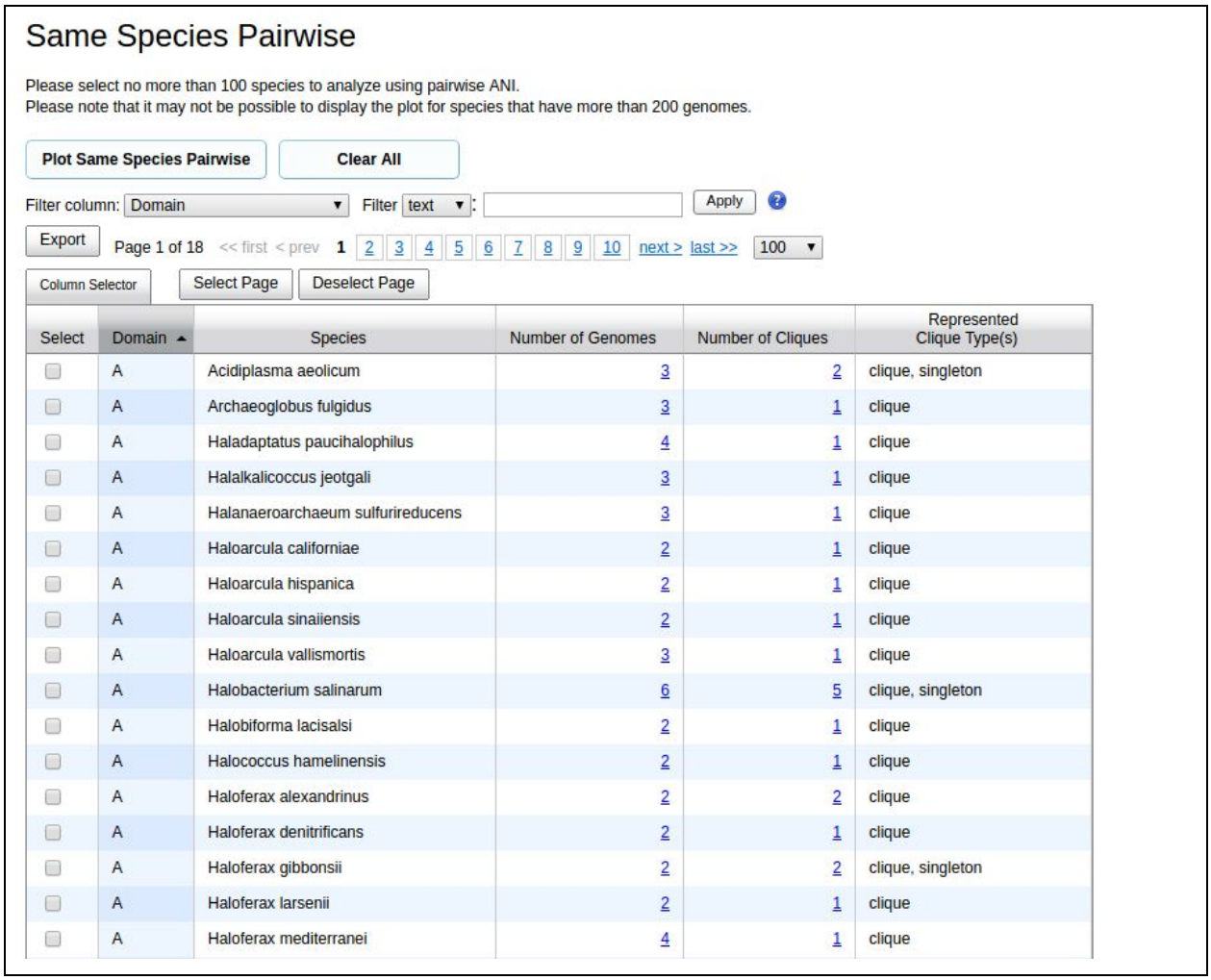
Export Page 1 of 1 << first < prev 1 next > last >> All

Column Selector

Genome1 ID	Genome1 Name	Genome2 ID	Genome2 Name	ANI1->2	ANI2->1	AF1->2	AF2->1	Total BBH	Precomputed ?
643348540	Desulfurococcus kamchatkensis 1221n	2513020047	Desulfurococcus fermentans Z-1312, DSM 16532	96.72	96.72	0.920	0.890	1305	Yes
643348540	Desulfurococcus kamchatkensis 1221n	649633040	Desulfurococcus mucosus 07/1, DSM 2162	73.26	73.26	0.610	0.610	830	Yes
643348540	Desulfurococcus kamchatkensis 1221n	2588253527	Euryarchaeota sp. genome BIN4	66.58	66.58	0.000	0.000	17	No
643348540	Desulfurococcus kamchatkensis 1221n	646564534	Ferroglobus placidus AEDII12DO, DSM 10642	73.26	73.64	0.010	0.000	26	No
646564534	Ferroglobus placidus AEDII12DO, DSM 10642	2513020047	Desulfurococcus fermentans Z-1312, DSM 16532	72.66	72.44	0.000	0.010	28	No
646564534	Ferroglobus placidus AEDII12DO, DSM 10642	643348540	Desulfurococcus kamchatkensis 1221n	73.83	73.26	0.000	0.010	28	No
646564534	Ferroglobus placidus AEDII12DO, DSM 10642	649633040	Desulfurococcus mucosus 07/1, DSM 2162	71.48	71.56	0.010	0.010	22	No

Figure 1b. Pairwise ANI - results

Same Species Plot



Same Species Pairwise

Please select no more than 100 species to analyze using pairwise ANI.
Please note that it may not be possible to display the plot for species that have more than 200 genomes.

Plot Same Species Pairwise | Clear All

Filter column: Domain | Filter text: | Apply

Export | Page 1 of 18 << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> | 100

Column Selector | Select Page | Deselect Page

Select	Domain	Species	Number of Genomes	Number of Cliques	Represented Clique Type(s)
	A	Acidiplasma aeolicum	3	2	clique, singleton
	A	Archaeoglobus fulgidus	3	1	clique
	A	Haladaptatus paucihalophilus	4	1	clique
	A	Halalkalicoccus jeotgali	3	1	clique
	A	Halanaeroarchaeum sulfurireducens	3	1	clique
	A	Haloarcula californiae	2	1	clique
	A	Haloarcula hispanica	2	1	clique
	A	Haloarcula sinaiiensis	2	1	clique
	A	Haloarcula vallismortis	3	1	clique
	A	Halobacterium salinarum	6	5	clique, singleton
	A	Halobiforma lacisalsi	2	1	clique
	A	Halococcus hamelinensis	2	1	clique
	A	Haloferax alexandrinus	2	2	clique
	A	Haloferax denitrificans	2	1	clique
	A	Haloferax gibbonsii	2	2	clique, singleton
	A	Haloferax larsenii	2	1	clique
	A	Haloferax mediterranei	4	1	clique

Figure 2. Same Species Plot - Species table

Selecting in the Species table under “Number of Genomes” for a species links to the “Genomes for Species” page (Fig. 2b). Selecting “Number of Cliques” for a species links to the “Cliques for Species” page (Fig. 2c).

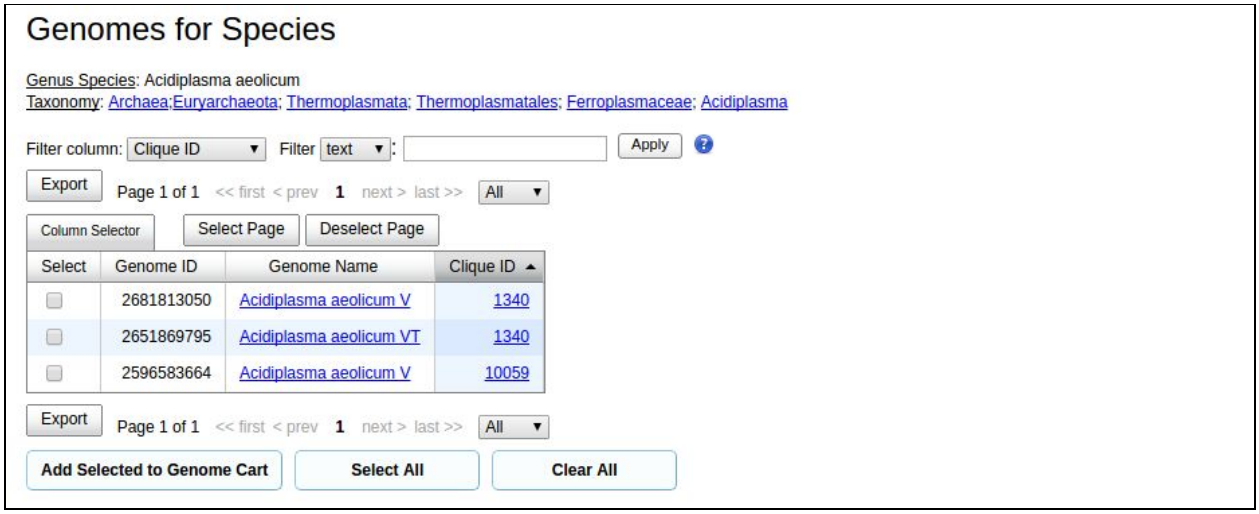


Figure 2b. Genomes for Species

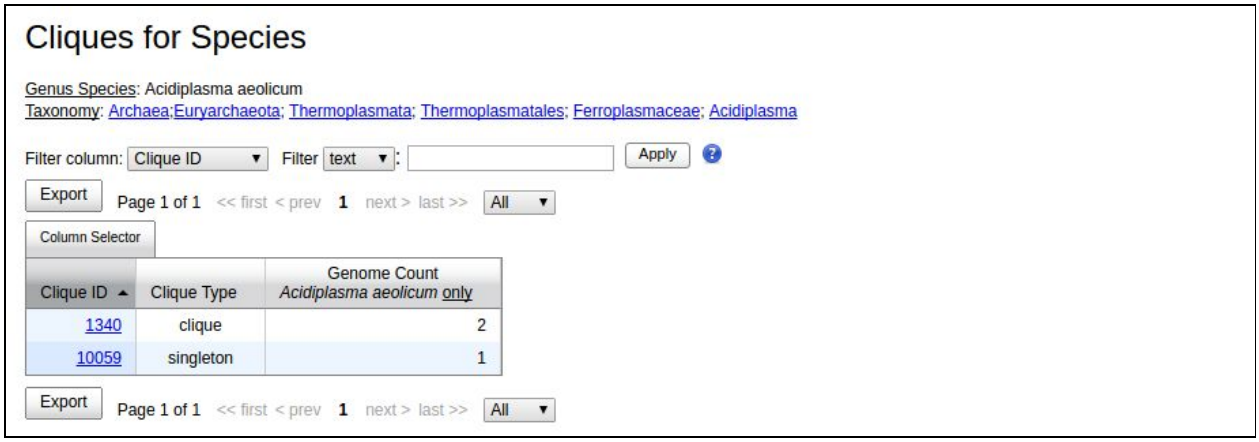


Figure 2c. Cliques for Species

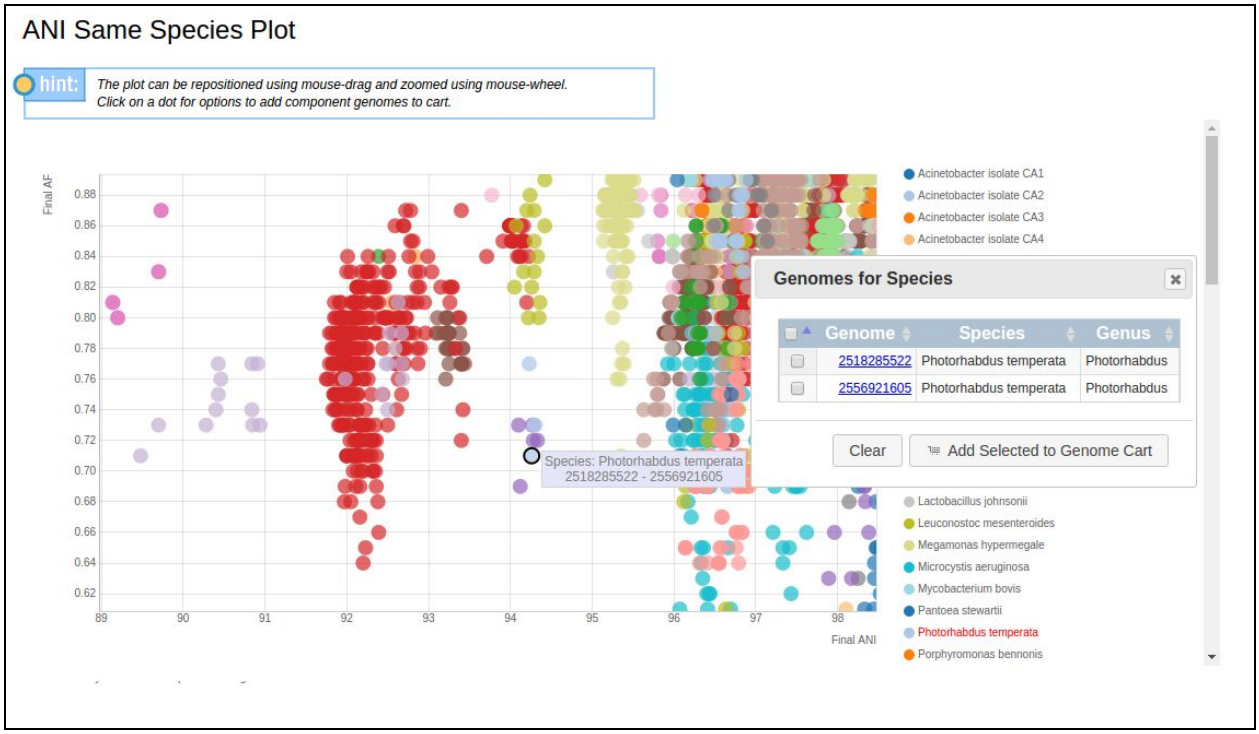
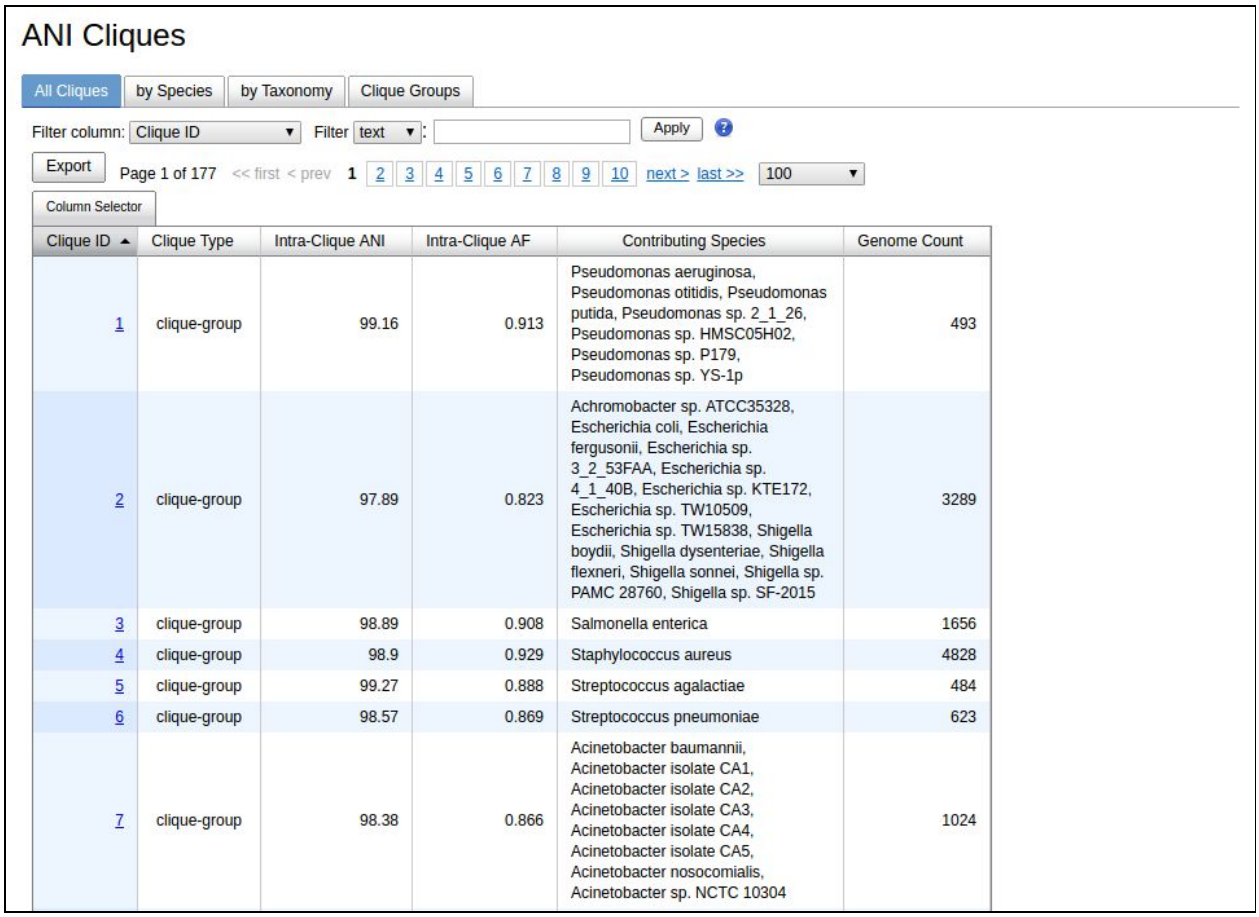


Figure 2d. Same Species Plot

Each dot in the Same Species Plot represents the final ANI vs. final AF for a genome pair present in a given species. Selecting a dot will bring up a dialog box with a table to which the 2 genomes represented in the dot will be added. These genomes can then be reviewed and added to a genome cart.

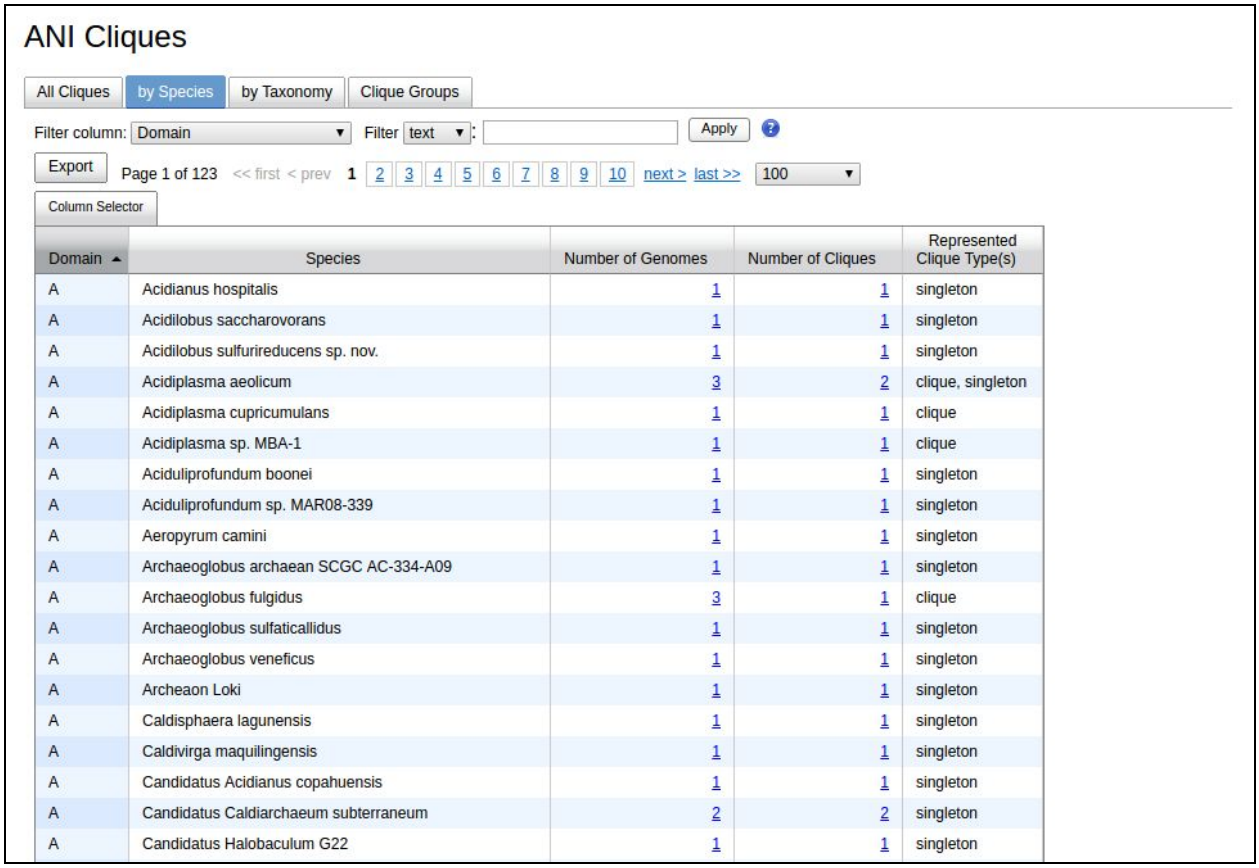
ANI Cliques



Clique ID	Clique Type	Intra-Clique ANI	Intra-Clique AF	Contributing Species	Genome Count
1	clique-group	99.16	0.913	Pseudomonas aeruginosa, Pseudomonas otitidis, Pseudomonas putida, Pseudomonas sp. 2_1_26, Pseudomonas sp. HMSC05H02, Pseudomonas sp. P179, Pseudomonas sp. YS-1p	493
2	clique-group	97.89	0.823	Achromobacter sp. ATCC35328, Escherichia coli, Escherichia fergusonii, Escherichia sp. 3_2_53FAA, Escherichia sp. 4_1_40B, Escherichia sp. KTE172, Escherichia sp. TW10509, Escherichia sp. TW15838, Shigella boydii, Shigella dysenteriae, Shigella flexneri, Shigella sonnei, Shigella sp. PAMC 28760, Shigella sp. SF-2015	3289
3	clique-group	98.89	0.908	Salmonella enterica	1656
4	clique-group	98.9	0.929	Staphylococcus aureus	4828
5	clique-group	99.27	0.888	Streptococcus agalactiae	484
6	clique-group	98.57	0.869	Streptococcus pneumoniae	623
7	clique-group	98.38	0.866	Acinetobacter baumannii, Acinetobacter isolate CA1, Acinetobacter isolate CA2, Acinetobacter isolate CA3, Acinetobacter isolate CA4, Acinetobacter isolate CA5, Acinetobacter nosocomialis, Acinetobacter sp. NCTC 10304	1024

Figure 3. ANI Cliques - by clique ID

Clicking on the clique id in the All Cliques tab will link to the details page for that clique (Fig. 4).



ANI Cliques

All Cliques | by Species | by Taxonomy | Clique Groups

Filter column: Domain Filter text: Apply

Export Page 1 of 123 << first < prev 1 2 3 4 5 6 7 8 9 10 next > last >> 100

Column Selector

Domain	Species	Number of Genomes	Number of Cliques	Represented Clique Type(s)
A	Acidianus hospitalis	1	1	singleton
A	Acidilobus saccharovorans	1	1	singleton
A	Acidilobus sulfurireducens sp. nov.	1	1	singleton
A	Acidiplasma aeolicum	3	2	clique, singleton
A	Acidiplasma cupricumulans	1	1	clique
A	Acidiplasma sp. MBA-1	1	1	clique
A	Aciduliprofundum boonei	1	1	singleton
A	Aciduliprofundum sp. MAR08-339	1	1	singleton
A	Aeropyrum camini	1	1	singleton
A	Archaeoglobus archaean SCGC AC-334-A09	1	1	singleton
A	Archaeoglobus fulgidus	3	1	clique
A	Archaeoglobus sulfaticallidus	1	1	singleton
A	Archaeoglobus veneficus	1	1	singleton
A	Archeaon Loki	1	1	singleton
A	Caldisphaera lagunensis	1	1	singleton
A	Caldivirga maquilingensis	1	1	singleton
A	Candidatus Acidianus copahuensis	1	1	singleton
A	Candidatus Caldiarchaeum subterraneum	2	2	singleton
A	Candidatus Halobaculum G22	1	1	singleton

Figure 3b. ANI Cliques by Species

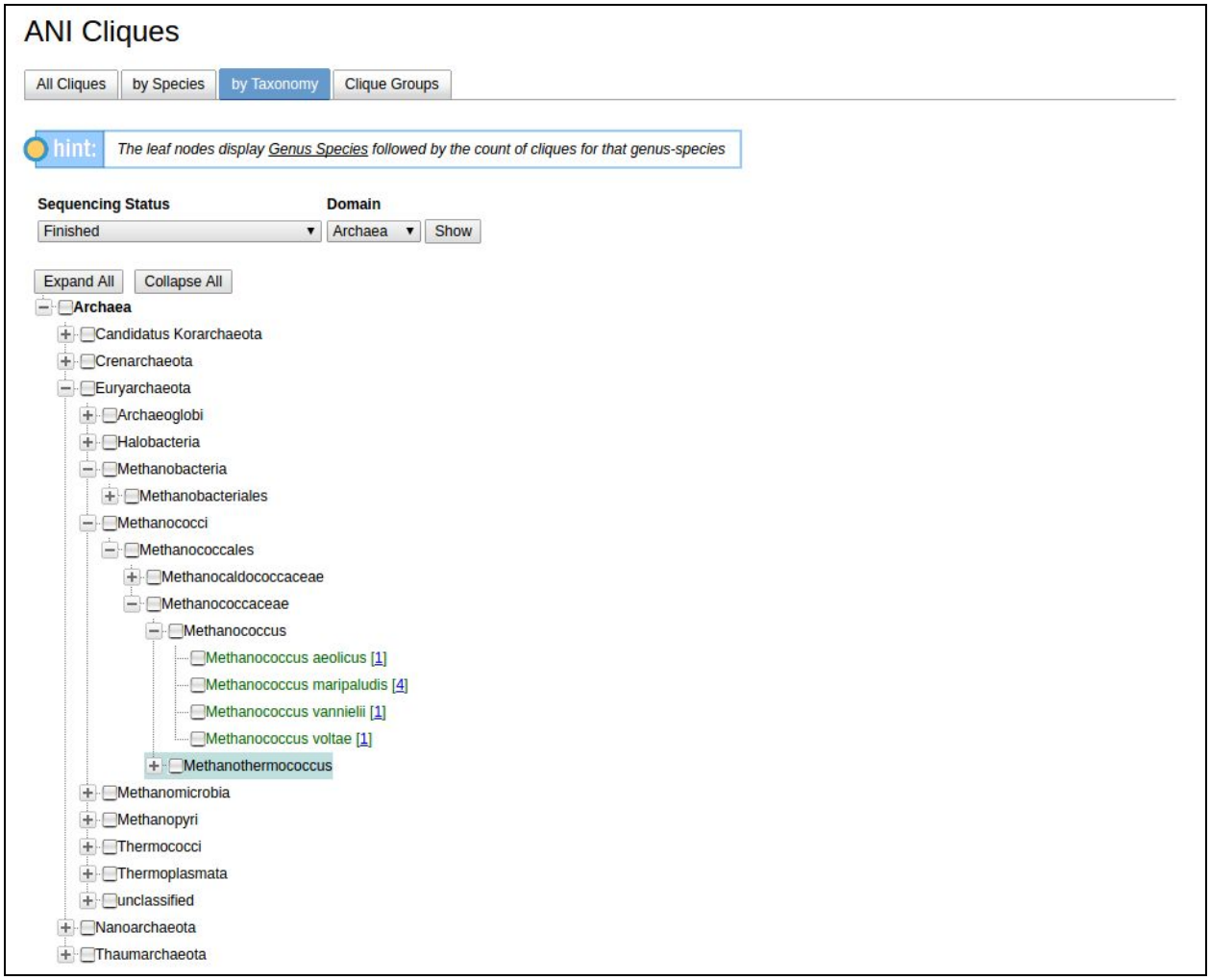


Figure 3c. ANI Cliques by Taxonomy

The leaf nodes in the Taxonomy tree display the Genus Species followed by the count of cliques for that genus-species. Clicking on the count links to the Genus Species detail page. The first tab of this page lists the cliques for this Species (Fig. 3d), while the second tab lists the genomes in that Species (Fig. 3f).

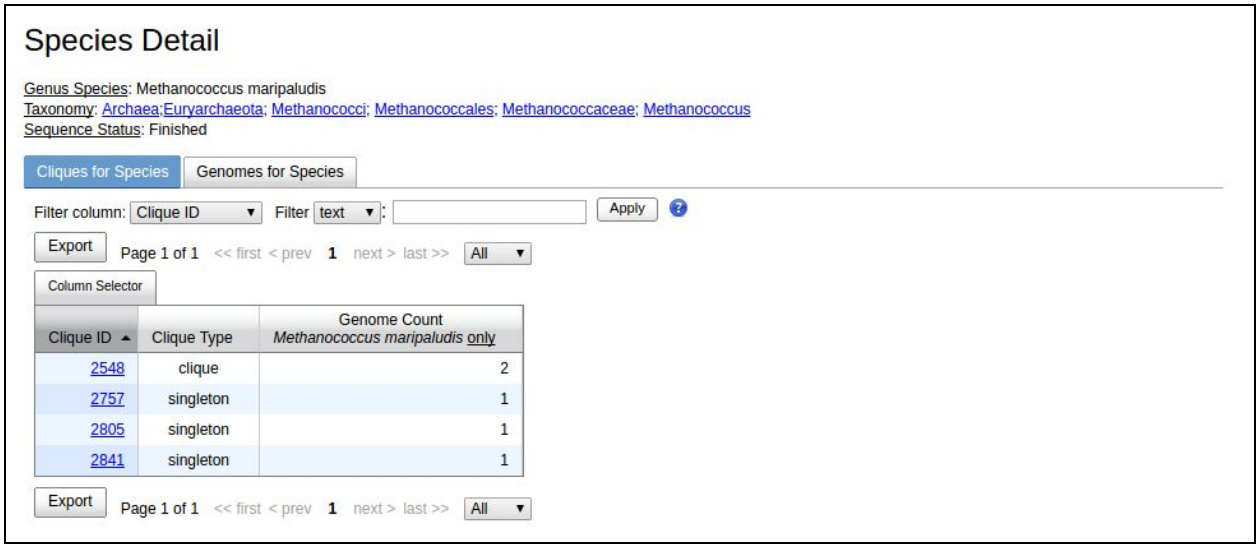


Figure 3d. Genus Species Detail page - Cliques for Species

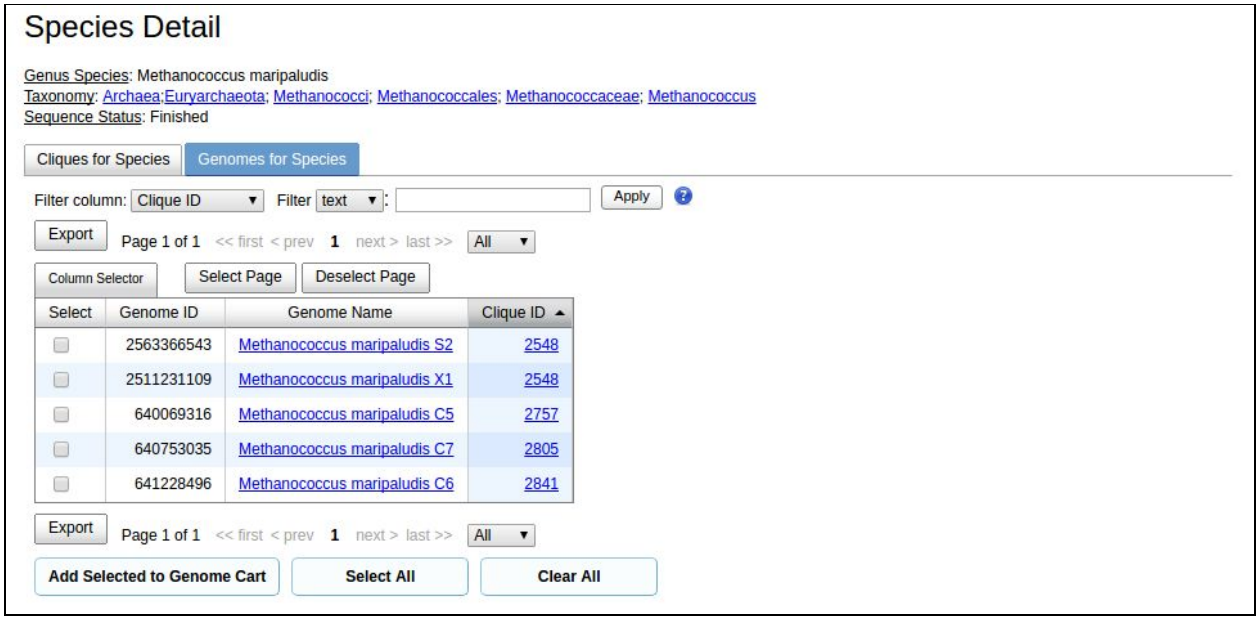


Figure 3f. Genus Species Detail page - Genomes for Species

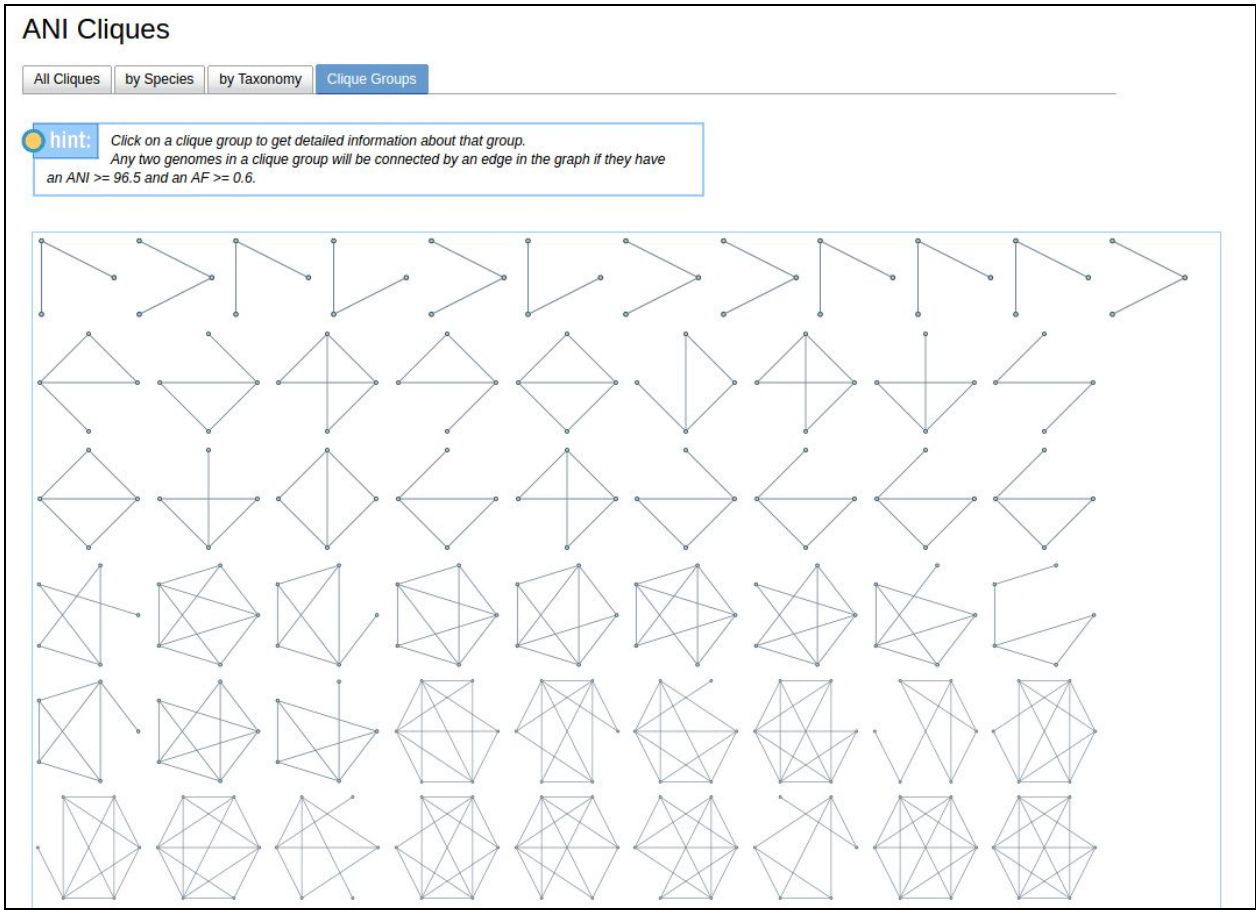


Figure 3g. ANI Cliques - Clique Groups

Any two genomes in a clique group will be connected by an edge in the graph if they have an ANI >= 96.5 and an AF >= 0.6. Clicking on the clique graph representation links to details page for that clique group (Fig. 4).

Clique Details

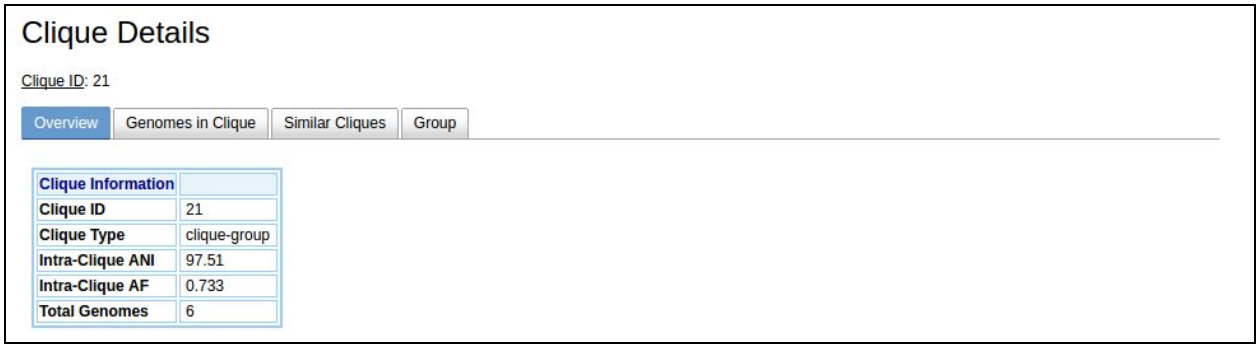


Figure 4. Clique Details - Overview

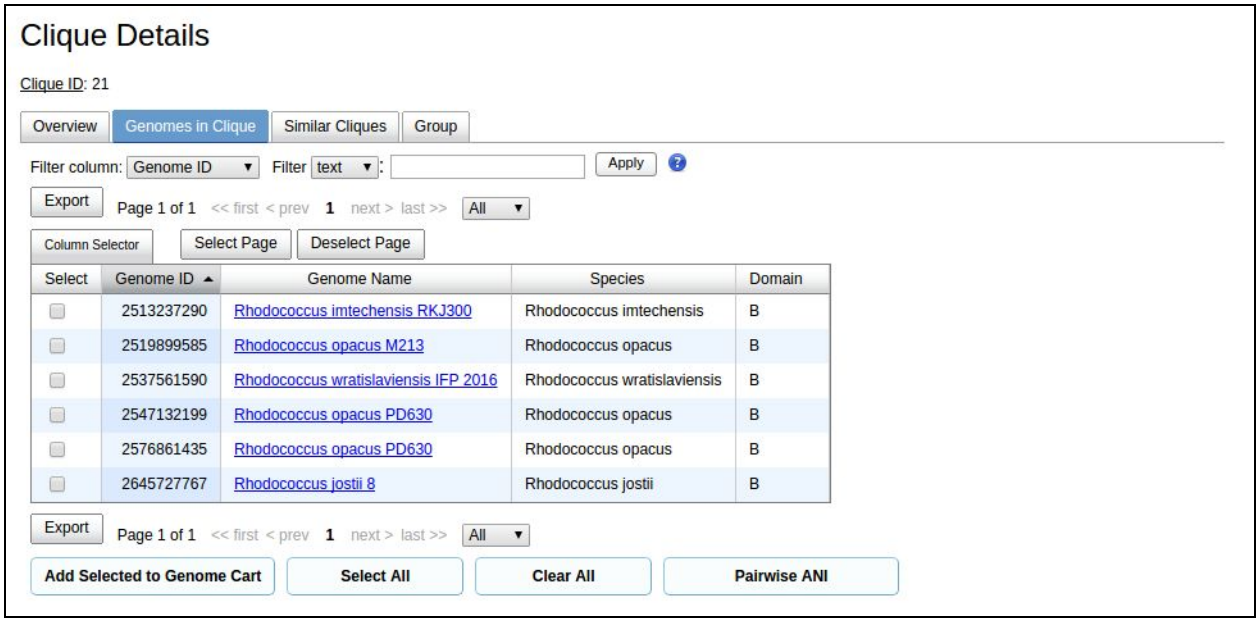


Figure 4b. Clique Details - Genomes in Clique

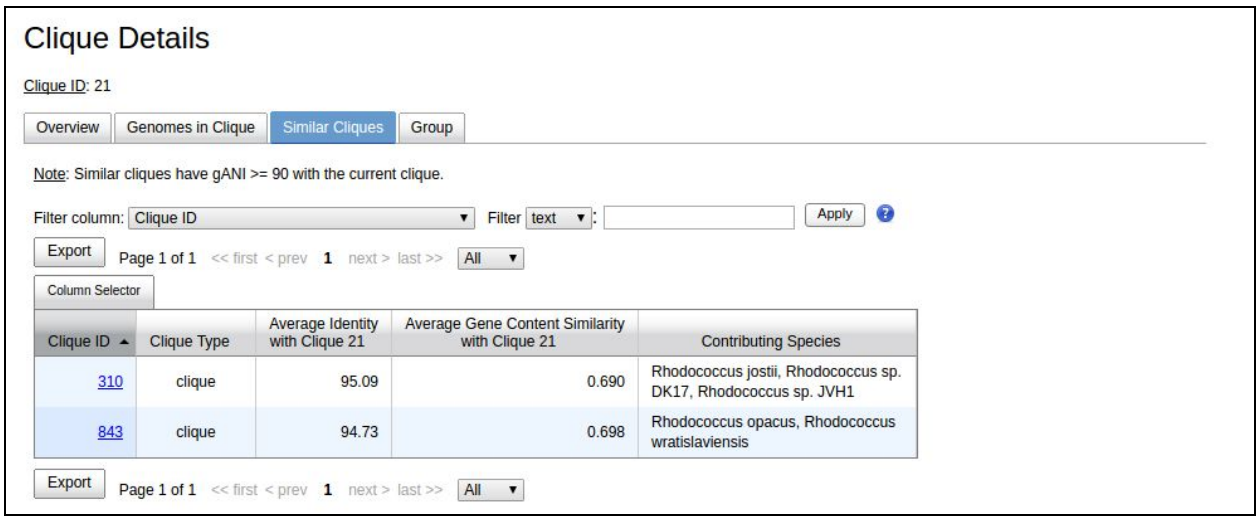


Figure 4c. Clique Details - Similar Cliques

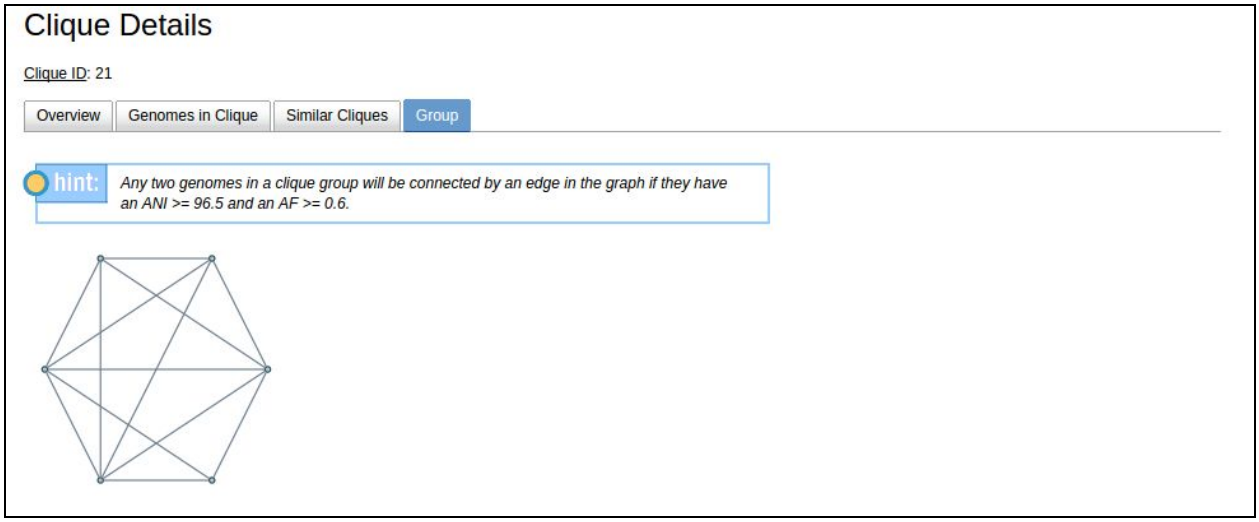


Figure 4d. Clique Details - Clique Group

Genome Detail Page

ANI information for a particular genome can be found on the genome detail page, located to the right of Genome Statistics.

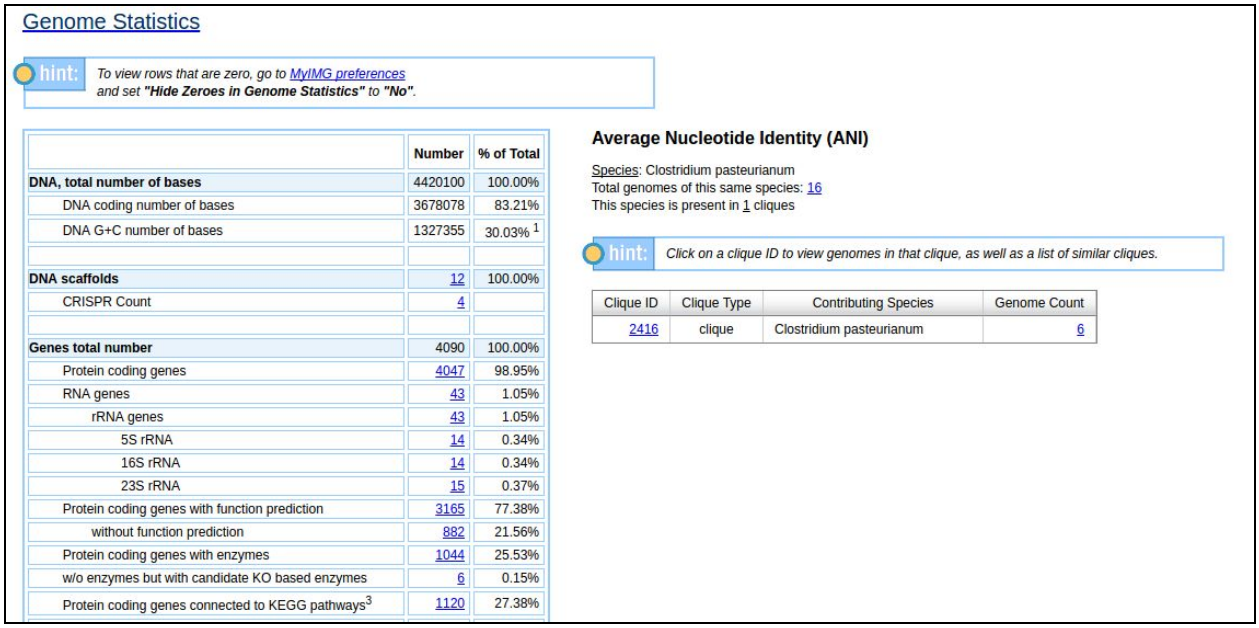


Figure 5. Genome Detail - ANI